DOCUMENT RESUME

ED 245 437 EC 162 714

AUTHOR Clarizio, Harvey F.; Mehrens, William A.
TITLE Identification and Programming for the Gifted: Some Limitations of Guilford's Structure of Intellect Model. Draft.
PUB DATE Apr 84
NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Development; Elementary Secondary Education; *Gifted; Intellectual Development; *Models
IDENTIFIERS *Structure of Intellect

ABSTRACT

The paper critically analyzes the research and promotional literature on J. Guilford's Structure of Intellect (SOI) model as it relates to the cognitive functioning of gifted students and evaluates the implications of this analysis for the counselor's role in the identification and programming of gifted students. The review is written from developmental and measurement bases and covers materials published by the SOI Institute, workshop materials from districts using SOI materials, and professional research literature. It is concluded that the SOI promotional literature promises far more than it has delivered and that severe psychometric limitations seriously limit the SOI model's usefulness. Limitations center on the lack of reliable assignment of test items to SOI categories, unreliable difference scores, and the lack of adequate validity evidence for differential treatment effects. (Author/CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED245437

Identification and Programming for the Gifted: Some Limitations of Guilford's Structure of Intellect Model

HARVEY F. CLARIZIO

WILLIAM A. MEHRENS

College of Education
Michigan State University

DRAFT
NOT TO BE CITED WITHOUT PERMISSION

Presented at the annual AERA meeting,
New Orleans, 1984

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Harvey F. Clarizio
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EC162714

Abstract

The purpose of this presentation is (a) to analyze critically the research and promotional literature on Guilford's Structure of Intellect (SOI) model as it relates to cognitive functioning of gifted students, and (b) to evaluate the implications of this analysis for the counselor's role in the identification and programming of gifted students. This critical evaluation is based on a review of promotional literature and available research, and is written from developmental and measurement bases. It is concluded that the promotional literature promises far more than it has delivered and that severe psychometric limitations seriously limits the SOI model's usefulness.

Identification and Programming for the Gifted:
Some Limitations of Guilford's Structure of Intellect Model

The field of gifted education has many unresolved issues. Among the most important is the identification and subsequent educational programming of gifted pupils. While the concept of intelligence has been broadened over the years, the newer concept remains to be adequately operationalized. Educational programs for the gifted vary in their objectives and in the instructional components they employ. The variability in educational programming is not unexpected in the absence of any commonly accepted operational definition of either intelligence or giftedness.

One attempt to define the concept of intelligence is reflected in J. P. Guilford's (1967) Structure of Intellect (SOI) model. In sharp contrast to the concept of unitary or general intelligence (g), Guilford, in his multifactor theory of intelligence, developed his well-known "mental block" consisting of three dimensions (operations, contents, and products) and 120 cells. Consistent with the multidimensional view of giftedness, SOI analysts seek individual patterns of relative strengths and weaknesses upon which to base an individualized educational program. Theoretically, examining cellular structure allows strengths and weaknesses in cognitive functioning to be identified in a highly specific manner. SOI profiles may be developed from several standardized intellectual tests (SOI Learning Ability Test, Stanford-Binet, WISC-R, Slosson). Curricular tasks can also be analyzed to identify which particular abilities are stimulated by selecting the appropriate curricular tasks.

The objectives of this paper are twofold:

a) To analyze critically the research and promotional literature on

Guilford's SOI model as it pertains to the cognitive functioning of gifted students; and b) To evaluate the implications of the analysis for the counselor's role in the identification and programming of gifted students.

Before we turn to the review, it is appropriate to state what specific materials we examined. Basically the review consisted of three parts. First, we read the most recent Product Catalogue to see what materials were published by the SOI Institute. From this catalogue, we reviewed: (a) Mary Meeker's textbook (1969) which explores and defines each of the SOI abilities, explains how to teach each of the abilities and describes how to make SOI profiles from the Stanford-Binet and Wechsler scales; (b) Collected Readings, Volume I which covers literature on the SOI for the years 1969-1975; (c) SOI Research Studies which contains synopses of empirical studies from 1981 to 1983; (d) the Technical Data Manual for the SOI-LA Basic Test for grades 2 through intermediate; (e) the SOI Form for the Gifted (G) which is a subset of the basic SOI-LA test and a group test used for screening potentially gifted students; (f) the SOI templates used to derive profiles for the Standford-Binet and the WISC-R; and (g) the Teacher's Guide for Using SOI Test Results, which explores the results as they relate to learning styles, learning problems and teaching strategies.

Secondly, we examined materials that we received from school districts that are using or have used SOI materials. These materials include workshop materials provided by national figures in the SOI movement as well as other training and inservice materials that school personnel received from the SOI Institute.

Lastly, our review covered the professional literature from journals, books, and presentations made at national conventions. While we would not want to make the claim that we left no stone unturned in our search for SOI materials, we clearly exerted more effort to identify and evaluate SOI

materials than would the typical school district for whom these materials are intended. Regretably, there is no single coherent source that provides the prospective user with adequate information upon which to base intelligent decisions regarding use of SOI materials. The evaluation of all of these materials was done from both a developmental framework and a measurement perspective.

The development framework will focus on whether the proposed and actual uses of the SOI model are congruent with what is known about cognitive development of children. From a developmental viewpoint, Guilford's theory fails to explain how people wind up having 120 different abilities. While developmentally oriented psychologists, like Piaget, note that not all individuals will achieve the highest levels (stages) of thinking, Guilford seems to state that intelligence subsumes all of the 120 abilities naturally.

The measurement perspective will focus on the whole notion of the SOI model from a psychometric point of view. How adequate are the norms for the various proposed uses? How reliable are the subtests, and more importantly, how reliable are the difference scores obtained from a series of SOI subtests? What is the evidence for the SOI model? Does the SOI model give us mathematical factors which do not necessarily have psychological/educational/instructional meaning? Is there compelling evidence for the usefulness of the model? (Even if a difference scores is reliable, it may not be valid.) What evidence exists that the matching of an individual's SOI profile to a particular instructional model enhances learning?

DEVELOPMENTAL CONSIDERATIONS

IMPLICATIONS FOR INSTRUCTION

Given the availability of SOI educational materials from age 6 months onward, advocates of the SOI model seem to convey the impression that all 120 factors are present from a young age even though they do not attempt to assess all "known" cells at any given age level. Yet, Piaget, a developmental psychologist who has given us the most comprehensive picture of cognitive development, notes that all abilities are not present in the young child. Instead the child develops intelligence over a period of several years. Even then, as many as 50 percent of adolescents and adults never reach the highest stage of formal operations. No allowance for developmental change is made in the SOI model. This approach largely ignores the changes in intellectual abilities that occur as the child increases in age. How many of the 120 abilities are present in the infant, in the preschooler, in the elementary school child, in the secondary student? No one addresses this issue systematically. Studies by other factor analytic theorists indicate that the number of abilities does vary with age. For example, research by Thurstone (1938) on the eight "primary abilities" identified in a population of adolescents and young adults showed that he could isolate only five comparable abilities among five and six year olds. As children grow older, particularly beyond the preadolescent years, there is evidence from factor analytic studies of increasing differentiation of intellectual ability (Ausubel and Sullivan, 1970). It is not at all clear from reading the SOI literature, how many abilities have been isolated at different ages. Studies using older and homogeneous subjects, (e.g., professionals) result in the emergence of numerous specific factors whereas studies using young and heterogeneous subjects generally provide evidence for a "g" factor (Ausubel & Sullivan, 1970). Thus it may be that the general

factor becomes differentiated during the adolescent years. Further research is needed to determine the extent to which Guilford's structure of intellect changes with age.

Not only is there no systematic discussion of developmental change with respect to the number of mental abilities (the quantitative aspects), but the SOI system does not allow for developmental changes in the qualitative aspects of intelligence. Piaget reported that thought is qualitatively different at each stage of development. Thought for the infant consists of sensorimotor coordinations, is perceptually dominated for the preschooler, is logical but concretely bound for the elementary school child, and is abstract and propositional for approximately half of secondary school students. While the SOI model purports to assess different kinds of intelligence, this model completely ignores fundamental qualitative differences that take place as the child develops.

Piagetian theory is also hierarchical. Indeed it is difficult to imagine that a child could grasp the nature of conservation without first developing an understanding of object permanence, i.e., an object can exist when out of one's visual field. In contrast to Piaget's hierarchical model, the SOI model is cubic (Sternberg, 1979).

Closely allied to the issue of changes in intellectual development are the issues associated with instructional strategies. Piagetian theory would indicate that just as the child's abilities vary with age, the educational implications also vary with developmental status. For instance, prior to the junior high period, it is especially important that the child be given concrete materials to manipulate in order to promote a clear and stable understanding of the world. Children must be given the opportunity to be active, to explore, to touch, to test, to find out and manipulate, because the kind of learning that Piaget proposed is a prerequisite for later verbal, abstract understanding. To understand classification the child must

have had concrete experience in grouping objects together in one or more dimensions; to analyze, she should have pulled them apart; to understand the ordered properties of number series, she must have arranged objects in some order. Gradually, these actions become internalized and carried out in the mind and the child's need for concrete aids diminishes. At this time, the adolescent becomes a more abstract learner and greater use can be made of verbal presentations. In brief, having recognized a fundamental change from concrete to abstract thought, as well as the hierachical nature of thought, developmentally based educators have adjusted their instructional methods accordingly. This developmental perspective is lacking in the SOI approach. This is no recommendation that methods be adjusted to the child's stage of cognitive development. One gets the impression that instructional strategies are as suitable at one period of development as they are at the next or at an earlier level. In their stress on intra-individual differences, SOI proponents have lost sight of nomothetic guidelines. Gifted children pass through the Piagetian stages of cognitive development at earlier ages than those of normal children, but there is no evidence to indicate that the sequence is different for gifted students. For the SOI faithful, the instructional methods do not necessarily or predictably vary with the child's stage of cognitive development. Instead, they vary instruction according to the bumps and valleys on the SOI patterning. This approach represents a form of psychometric phrenology except that the diagnosis substitutes the feeling of bumps on the child's SOI profile instead of the child's head.

THE MYTHICAL CONCEPTION OF NORMALITY

Over the years, attempts have been made to use profile analysis of intelligence tests such as the Wechsler Scales for diagnostic and educationally prescriptive purposes with various groups of exceptional children (Clarizio and Veres, 1984). For example, early work on the identification of learning disability was based on the implicit assumption that children who are having difficulty succeeding in school, which can be traced neither to a physical handicap nor to mental retardation, must be deficient in some component of cognitive ability. This assumption, coupled with the notion that intelligence tests assess the structure of cognitive processing, led to the search for characteristic patterns in the intelligence test profiles of these children. The widespread use of the Wechsler Intelligence Scales (WISC/WISC-R) and the availability of their subtest scores led naturally to hypotheses regarding connections between particular patterns of scores and learning disabilities. Uneven function hypotheses have also appeared in the emotional impairment literature. The rationale behind these formulations is less clear than in the learning disabilities area. To some extent they were guided by the expectation that psychosocial deficits would have a more profound effect on verbal than on non-verbal abilities. In both areas, the early research results seemed promising. Although early work in these areas was somewhat encouraging in that patterns that seemed to characterize the subtest scores of learning disabled and emotionally disturbed children were identified, later research showed these efforts to be misguided (Clarizio and Veres, 1984; Clarizio and Bernard, 1982; Kaufman, 1981.).

Until recently, it was assumed that the "normal" child had a uniform distribution of cognitive abilities. After all, the mean score on each

cognitive scale is, in fact, its norm. It was not until the mid-1970's that the extent of variation in the profiles of normal children was realized. Up to that time, most research proceeded with neither control groups nor any other form of base rate information. It became evident, to some, that the rather variant subtest patterns that had been thought to be indicative of educational disabilities or emotional disturbance were present in the profiles of many children who were experiencing no educational difficulties. In short, variablity as evidenced by peaks (strengths) and valleys (weaknesses) on profile analysis of cognitive measures is a common attribute among normal and handicapped groups. Moreover, available evidence suggests that such variability is even greater among gifted children (Sattler, 1982). Yet, SOI advocates, who also favor pattern analysis, seem to regard variability particularly below the child's mean in cognitive functioning as a condition in need of special treatment or remediation even though these children were selected and tested in the absence of any particular concern on the part of their teachers, and not because of poor performance in the classroom. As diagnosticians, it is understandable but regretable to err by attributing pathological significance to various measures of test scatter among handicapped children in that the two conditions (intratest variability and handicaps) covaried. It is inexcusable for us to do so with gifted children who are functioning well in school. Consistent with the interpretations of valleys on the profiles of learning disabled children as deviant, the valleys on the profiles of gifted children are also interpreted as being unusual if not pathological. This assumption is indeed questionable in light of recent evidence that variability in cognitive patterning is far more typical than previously believed by psychologists and that this variability is even greater among gifted children than among those with lesser abilities. We must remember that these gifted children were not

identified because of unsatisfactory performance in the classroom.

The finding of intratest variability among gifted youngsters is not restricted to traditional individual intelligence tests. Uneven test performance also holds with respect to SOI scales. Yet despite the awareness by some SOI followers that "for children of very superior IQ status uneven performance is the rule rather than the exception" on SOI measures for both boys and girls (Millichamp, no date), the majority of the SOI literature suggests that "irregularities" within each profile "should be the basis for remediation." The researcher might be thinking that a little extra educational attention, even if it is directed at imaginary strengths and weaknesses, certainly can not hurt the gifted student. Therefore, why be so critical of profile analysis? Indeed, what might appear innocuous, or perhaps even beneficial, is the SOI follower's emphasis on peaks or strengths. It is, of course, appropriate to provide enrichment in areas of a child's strengths. However, we must first be sure that the peaks truly represent strengths and are not minor flucuations or chance deviations in a child's performance. Even if the strengths do represent reliable differences from the child's overall performance, the child may not necessarily be in need of enrichment. Statistical differences are not necessarily educationally significant. Given the finite amount of resources made available to the gifted in our society, we can ill afford to concentrate our time, energies, and dollars on conditions that are not in need of special attention. We also run the risk of making the gifted child feel that she is weak or strong in given abilities when this is not the case. In short, the unwarranted assumption about the evenness of cognitive patterning could well result in a large number of false positives i.e., children who we say have a changeworthy condition but who in reality are normal.

PSYCHOMETRIC CONSIDERATIONS OF THE SOI MEASURERS

For any test to be useful, it must meet certain basic psychometric standards with respect to such factors as reliability, validity, and adequacy of normative data. The APA/AERA/NCME Standards for Educational & Psychological Tests (1974) presents standards for test use as well as for test manuals and is intended to guide both test developers and test users. In general, the SOI-LA does not do as well as most commercially sold tests with respect to meeting the standards stated in that document. We will briefly discuss how the SOI-LA fares on four of the topics covered in the Standards: Norms, Reliability, Validity, and Aids to interpretation.

Norms

The Technical Data Manual (1981) reports that the SOI-LA was re-normed in 1980. There were only six sites in the norm group. Three of these were located in California. The others were in Indiana, Oklahoma, and Texas. There is no information presented as to why these sites were chosen, or how the children were selected within these sites. There were a total of 2014 students tested (1015 males, 999 females) in grades 2 through 6. The manual states that the norms on this SOI-LA test apply to nine of the subtests on the Gifted Screening Test although we are not told how many, if any, in the sample were supposedly gifted. No reason is given as to why they do not apply to the other three subtests used for form G (NSI, DFU, & DMU). Compare this brief description with Standard D2 in the afore mentioned book: "Norms presented in the test manual should refer to defined and clearly described populations. These populations should be the groups with whom users of the test will ordinarily wish to compare the persons tested. Essential" (APA, 1974, p.20). Further, it is not at all clear why the

profile on form G (copyrighted in 1975) has norms for intermediate and adult levels when the norm sample only included students up through grade 6.

Reliability:

The reliability data presented in the Technical Data Manual were based on a test/retest procedure with a two-to-four week interval. Alternate form correlations are also reported. Reliability data are presented for each of 26 subtests; for each of the 14 general ability tests, and for the overall scale for each grade from grade 2 through grade 6. The manual also reports what is incorrectly called the Standard Error of Estimate which was based on the test/retest correlations. The more standard statistic to report would be the standard error of measurement, and indeed, a little arithmetic shows that is what is reported--it is just misnamed in the manual! The manual reports nine sets of such statistics for each grade based on nine cells in the norm groups: Form (Form A, Form B, Total) crossed with sex (Males, Females, and Totals). The authors underline what they refer to as the "preferred set(s) of statistics" and those are the ones we looked at. The sample sizes for the cells ranged from 71 to 474 (for both sexes and both forms in grade 4).

In general, the reliabilities (test/retest correlations) of the subtests are quite low. For example, the reliabilities for the DFU subtest range from .36 to .60. For the DSR subtest, the test/retest correlations ranged from .27 to .55: for CFC from .15 to .32. Of course, not all were this low, but even for the subtest that appeared most reliable (CFU), the correlations ranged only from .60 to .77. These are hardly high enough to allow much confidence in the scores, and certainly would lead to unreliable difference scores. Also, recall that these coefficients are based on a test/retest at a two-to-four week interval. Certainly we would expect much lower correlations for longer intervals. Data only a few weeks old would almost certainly not be reliable enough for decision making.

The reliabilities of the general ability composites were, as expected, a bit higher. However, even for these composites, the majority of the correlations appeared to range from the low .50s to the mid .70s. Some of the correlations ranged into the .20s. The very highest correlation was for the Symbolic test (.90 for both sexes for form A in grade 6). But for males for Form B in grade 3 the Symbolic reliability was only .54.

The reliabilities for the overall scores on the SOI-LA ranged from .63 to .92. This is lower than one might expect for a test as long as this one is.

The manual does <u>not</u> report the reliabilities of the difference scores. However, because the intercorrelations are reported, they can be computed. We computed two such reliabilities: one for subtest differences and one for composite differences. We used the first intercorrelation given in the matrix for grade 2 in each case. For the difference between DMU and DFU for grade 2 we obtained a reliability of .42. For the difference between the composites of memory and cognition for grade 2, we obtained a reliability of .355. (One might think that composites should have more reliable difference scores than the subtests because they are more reliable. However, this is countered, at least in the example above, by higher intercorrelations between the composites which lower the difference score reliabilities.)

Given the generally low reliabilities of the difference scores and the total number of difference scores an individual teacher might look at given 20 to 30 children each taking all the subtests of the SOI-LA, one would surely expect to see a large number of differences that would most reasonably be interpreted as due to chance alone. We find it interesting (and deplorable) that neither the manual nor the computer printout of an individual's profile mention the problem of unreliable difference scores and how one should be cautious when making inferences from observed

differences--particularly when so many difference scores are being compared. It is not even clear how large or small an observed difference Meeker is willing to interpret as a true difference.

Comparing many different scores on a profile is analogous to running multiple t-tests. Some are almost sure to be "significant" due to chance alone. Of course there are ways to control for this (Silverstein, 1982) but this manual does not even talk about the basic notion of difference score reliability let alone considering some of the problems and solutions when inferring an observed difference is, or is not, a true difference.

One might compare what was done with the SOI-LA with the Standards where it is stated that: "The test manual or research report should present evidence of reliability, including estimates of the standard error of measurement, that permits the reader to judge whether scores are sufficiently dependable for the intended uses of the test" (APA, 1974, p.50). If one is to use the test for profile interpretation, the reliabilities and the standard errors of the difference scores should be reported.

Reliability of Template Scores

In addition to building tests based directly on Guilford's SOI model (SOI-LA, SFG), Meeker (1969) has devised a method that uses templates to analyze commonly administered intellegence tests. Let us consider some of the more serious drawbacks associated with SOI profile analyses with the Stanford-Binet L-M, a scale which is said to sample more of Guilford's cells than the WISC-R, the Slosson, or the Detroit Test of Learning Abilities. To determine strengths and weaknesses on the profile, Meeker uses a criterion of four months of mental from the child's mental age as a criterion of significant deviance (Millichamp, no date). To begin with, a deviation of 4 months of mental age represents one-sixth (16 percent) of the child's total mental development for a two-year old; however, at age 10, a deviation of four months represents only 3 percent of the child's development. Another

complicating factor with the use of a four month criterion is that this set standard does not control for the variability in IQ distributions that has been found to exist at different age levels. Sattler's standard deviation approach is a sounder method than Meeker's four month criterion. Another problem is that psychologists have difficulty classifying Stanford-Binet items in accordance with SOI categories. In one study (Dyer, Neigler, and and Milholland, 1975), nineteen school psychologists attempted to assign 142 items on the Stanford-Binet, L-M, to the five operations categories of Guilford's SOI model following charts prepared for this purpose by Meeker. On the average, one rater agreed with another on about half of the items and their modal assignments agreed with Meeker's assignments on only 57 percent (N=81) of the items. These levels of agreement are not considered high enough to justify classifying Binet items in accordance with SOI categories. Further, many of the items of the Binet as rescored have multiple classifications for SOI operations. For these items, success or failure must be assigned to each of the abilities involved, in keeping with the Binet procedures. Also, as can be noted from the templates, the number of measures for each SOI ability varies from one age level to another. Finally, it should be noted that the Binet measures some SOI abilities more thoroughly than others at a given age. To establish the specific effect of these inconsistencies on SOI ability scores would require item analysis of each child's Binet profile (Millichamp, no date). Given the above shortcomings, we cannot recommend profile analysis of Binet based on SOI categories.

Validity:

The Technical Data Manual says nothing directly about validity. One could made some very weak inferences about the degree of construct validity by correcting the intercorrelations for attenuation but the authors of the

manual did not even do this much. We did correct the intercorrelations between cognition and memory for attenuation for grades 2 and 6. We obtained corrected correlations of .68 and .85. Thus, we could say that the data support the SOI model in the trivial sense that the corrected correlations are not perfect. On the other hand, .85 seems quite high. Most would suggest correlations as high as .68 or .85 indicate that a general factor exists between those two composites.

Some inferences can be made about the degree of validity of the SOI-LA for various purposes by reading the reports in the documents entitled Collected Readings (SOI Institute, no date) and SOI Research Studies (SOI Institute, no date). However, it is hard to know just what should be inferred from these documents. For example the document entitled Collected Readings-Volume I containes 38 readings. Of these, 23 were authored or coauthored by Meeker. Only 19 of these 38 readings were what we would call research, and of these, only four were published in a professional journal or book.

The SOI Research Studies document is a collection of studies of the SOI. These studies are classified into three groups: Level I--Studies focusing on reliability and validity of SOI assessment instruments; Level II--Studies focusing on the efficacy of SOI training-measuring growth in abilities that results from SOI training; and Level III--Studies focusing on the carry-over of SOI training to improved performance in school and work.

There were eleven Level I studies, eleven Level II studies, and seven Level III studies. A reasonably typical example of a Level I study is one that had as a purpose to determine if certain SOI learning abilities are related to reading achievement. Another Level I study investigated whether SOI learning abilities were related to measures of self-concept. A typical Level II study sought to determine whether SOI training produces an increase in learning abilities. An example of a Level III study was to determine

whether SOI training would lead to improvement in arithmethic achievement. The studies described in the document are all those collected from 1981 through 1983. No criteria are stated for the studies being included in the booklet, except that they are empirically based. It is reasonable to assume the studies are not an unbiased selection from all relevant studies. Of the 29 total studies in this document, only 2 appear to have been published in professional journals. From reading the short (one or two page) summaries of the studies it is usually apparent why they would not have been published. Research flaws are reasonably numerous.

Of the 11 Level 1 studies, none of them actually relate to the reliability of profile analysis or the increasing of differential predictive validity, although there are some studies showing the SOI-LA predicts school achievement, as would any aptitude test. One of the Level I studies shows that "various kinds of memory appear to be independent of one another." It is not stated how this conclusion was arrived at. Were the correlations among the memory tests not statistically significant? Were they less than one? Were corrections made for attenuation? Another Level I study shows that SOI-LA subtest scores can be used to "predict teacher perceptions beyond a chance level." This seems a fairly trivial and obvious finding. Two other Level I studies show a trend in the direction the researchers expected, but no statistical significance. One other study does not do any cross-validation of a regression equation.

Of the eleven Level II studies, five did not have a control group so one could not differentiate such things as practice effects on the test from training effects. One of the others had no statistical analysis, one had a trend in the 'right' direction, but a nonsignificant statistical test. Several of the Level II studies did show that one can raise scores on the SOI subareas by teaching those subareas. This seems much like raising the

score on an eye exam by teaching the letters on the eye chart. Any time one raises the score on an aptitude test by teaching toward the test one should consider whether the test has intrinsic or extrinsic predictive validity (Gulliksen, 1950). If the test has intrinsic predictive validity, this means it predicts because the skills measured on the test are important skills to have in order to do well on the criterion measure. If a test has extrinsic validity for predicting achievement this means that the skills measured on the predictor are not in and of themselves important for success on the criterion, but rather, that both measures are related to same third variable. Although it is a goal worth striving for in test development, most tests do not have intrinsic validity. Training for improvement on most aptitude tests does no more for achievement than training on an eye chart does for vision.

All of this is recognized at least intuitivily by those individuals who conduct studies like those categorized as Level III studies. Do these studies indicate that the SOI-LA tests are intrinsically valid? None of the seven Level III studies were published, although they do seem to provide a little supporting evidence that training on the SOI carries over to improved performance in school work. However, research limitations of the studies suggest that the inferences must be made with considerable caution. For example, three of the seven studies had no control group and one of the other studies had only a "quasi-control" group. One of the remaining three studies did not show any significant differences between experimental and control groups on the achievement variable. This leaves us with only two studies from which to make inferences regarding the carry-over of SOI training the school achievement. In our view, the case for intrinsic validity of the SOI-LA has not been substantiated.

Research studies in published journals not mentioned in the SOI literature are less positive than those that are mentioned. One concurrent

validity study of the SOI-Form G indicates that the multiple correlation between the WISC-R Full Scale IQ and the weighted subtests of the SOI-Form G was only .337, explaining about 11% of the total variance (Stenson, 1982). Another study, investigating the efficacy of the SOI as an identification tool for the gifted, indicated that the SOI-LA Form G does not show a strong relationship to success in the gifted program. The authors suggested that its correlations with academic variables may not be significant enough to warrant its use as an identification tool, particularly in view of the cost per pupil involved in administering the instrument. In regression equations, the SOI-LA Form G contributed only minimally to the predictions for an annual evaluation of academically gifed students or other academic variables (O'Tuel, Ward, & Rawl, 1983).

A study by Pearce (1983) compared the WISC-R, Raven's Standard Progressive Matrices and the SOI-Screening Form for gifted. She found the SOI abilities had varying relationships with intellectual functioning as measured by the WISC-R and concluded that using the SOI-G for either screening or identification purposes should be attemped cautiously, if at all.

Other reviewers of the SOI model and the SOI-LA test also have been somewhat cautions and/or critical. Eysenck (1967; 1973) and Cattell (1971) believe that the SOI model could be considerably simplified by combining factors that overlap. Horn and Knapp (1973) criticize the subjective element in the choice and rotation of factors followed by Guilford. This is particularly troublesome where the factors are rotated to conform with the specified SOI model. McNemar (1964) also believes the SOI system is far too elaborate. He objects to the "fractionization and fragmentation of ability into more and more factors of less and less importance" (1964, p. 872).

Vernon (1979) objects to the fact that Guilford's approach lacks good evidence to prove that the batteries "measure recognizably distinct abilities in daily life" (p. 60). Vernon (1965) has suggested that "a large proportion of Guilford's numerous factors of intellect have failed to show any external validity which could not be accounted for by their g, v, and space content" (1965, p. 726). Jensen (1980, p. 225) in discussing the SOI model comments as follows: "It has not been satisfactorily demonstrated that each of the 120 cells represents a different ability, except in the trivial sense, that any tests that correlated with all the other tests less than perfectly, after correction for attenuation, represent separate ability factors. The method of factor analysis allows almost infinite subdivision of abilities if one wishes to identify factors that reliably account for almost vanishingly small percentages of variance among all ability tests."

Cronbach (1970) also is critical of Guilford's model. "When Guilford says that 82 factors of the 120 have been 'discovered' or 'demonstrated', he provokes disagreement. . . Clearly, Guilford acknowledges a factor as worthy of attention long before other experts are ready to do so" (Cronbach, 1970, p. 339). As Cronbach points out: "The burden of proof is always on the person who advocates recognition of a new factor. He ought to show that the new factor accounts for an appreciable amount of variance that other factors cannot account for, and that the newly proposed factor has some interesting correlates with non-test behavior" (p. 342). We, and the reviewers mentioned here, do not feel comfortable that such proof exists. Of course, Meeker disagrees. As she argues in the Teacher's Guide, "Educators in their search for the best for students are often confused when some 'expert' in statistics maintains factor analysis (and the SI model) is not 'valid.' Such statements are not only false, but indicate as well that the 'expert' has not followed SI theory nor read about SOI practices and research" (Meeker, 1981, p. 13). Her defense is weak. The references Meeker

cites in the paragraph from which we quoted do not effectively counter the criticisms. Some 'experts' have followed SI theory and SOI practices and have remained both doubtful and critical.

Aids To Interpretation

As we have discussed, the norms, the reliability, and the validity of the SOI-LA have all been subject to some criticism. Do these criticisms have any relevance for someone interpreting the test? Are there statements in the SOI-LA materials that relate to how data from the test should be interpreted? The APA Standards state that: "The test, the manual, the record forms, and other accompanying material should help users make correct interpretations of the test results and should warn against common misuses" (1974, p. 13). Actually, there is little evidence that any information exists that would warn users against common misuses. In fact, we would submit that common misuses are indeed encouraged. Recall the quote before from Meeker regarding the criticisms of the SOI model. Basically she discounts the criticisms and implies that one would indeed be quite foolish to put any stock in those silly statisticians who misunderstand her work.

What "aids" are available, and what do they say? Actually, there are a fair number of short papers printed by the SOI Institute that presume to be aids to educators. One such 6 page paper is titled "Using cognitive abilities as the diagnostic basis for curriculum planning." In this paper Meeker suggests that by using the templates for scoring the Stanford-Binet, "individual curriculum planning can be done" (Meeker, no date, p. 1.) In a section under Convergent Production in that paper she suggests that: "Usually the superior student reproduces exactly what the teacher wants and is often, therefore, misjudged as gifted" (p. 3). (italics added). We leave it to you to determine why it is a misjudgment to consider a superior student gifted. On page 5 of the same document, Meeker states that: "Good

patterns in divergent production usually are significantly related to personality patterns in which good self-concepts exist" (p. 5). This is an interesting statement because in the SOI Research Studies publication, the only study that relates SOI abilities to self concept found that "the total self-concept yield on the Piers-Harris' scale was found to correlate significantly with the SOI operations Memory, Evaluation, Cognition, and Convergent Production. Little relationship was found between Total Self-Concept Scores and Divergent Production. . .. Low correlation between Divergent Production and self-concept support the idea that Divergent Production. . . is not influenced by self-concept" (Study 12/81-6 in the SOI Research Studies).

In the publication Using SOI Test Results: A Teacher's Guide (Meeker, Copyright1979, revised 4/81) many statements are made about what various subtest scores mean, and many suggestions are offered to teachers. We would consider many ill-founded. Nowhere in this guide are any cautions given regarding the unreliability of the subtests, or the limited validity of the SOI. Consider the following quote regarding the CSR--Comprehension of abstract relations--subtest. "CSR is the most difficult task on the SOI-LA test, even though there are only eight items. [a non-sequitor] We use CSR as an indicator for youngsters who say they want to go into the computer sciences, theoretical math and physics. . .. We interpret CSR as an ability to discover abstract relationships" (1981, p. 5). The Technical Data Manual shows that the CSR has reliabilities of .26 for both sexes, both forms for grade 2, .05 for males for both forms in grade 3, .28 for both sexes for form A in grade 4, and .50 for both sexes both forms for grades 5 and 6. We submit that most individuals who knew a bit about reliability would be somewhat cautious about making inferences about individuals from an 8 item test with reliabilities this low. Because most teachers will not

read the technical manual, it appears very unfortunate, and perhaps even unprofessional, not to mention any cautions about interpretations of such scores in the Teachers Guide.

Finally, we should mention the individual profiles which can be obtained as computer printouts. These profiles do not contain any band interpretation, there is no indication regarding how to interpret a difference score, and there is no statement warning a user regarding the unreliability of difference socres.

In conclusion, we believe the aids to interpretation are extremely weak with respect to the APA standard quoted earlier that the accompanying material should warn users against common misuses.

CONCLUSIONS AND IMPLICATIONS

Has the intellectual pie been sliced too thin? Of interest from both a theoretical and practical standpoint are the efforts by SOI advocates to continually reduce the number of factors for which they test. Initially, 120 factors were identified. Then Meeker identified 26 factors as most critical to school learning and devised the SOI-LA scale to measure those "basics" in education. A screening form has now been devised that tests for 12 factors. Given this progression of events, it is conceivable that SOI researchers will rediscover the "g" factor by the year 2000. It does seem that Guilford's "mental block" does indeed partition intelligence into too many pieces to have educational relevance. Indeed, it is not at all clear that SOI tests will yield higher correlations than traditional "g" factor tests with current or future school achievement. McNemar's (1965) conclusion that a large number of Guilford's factors failed to show external validity to any greater extent than could be accounted for by the "g" factor or by group factors seems as true today as it was when he made it almost two decades ago.

We offer the following closing comments.

1. The promotional literature promises more than the use of the SOI model itself can deliver. Certainly, the research basis for the promotional literature is sketchy.

2. Specialists in the gifted must take into consideration not only quantitative variations but also qualitative variations in assessing cognitive development.

3. SOI advocates assume that variablity, both below and above the child's mean, is indicative of a deficiency or a strength and should be remediated or enriched (Kester, 1982). Yet, recent studies have shown that variability is more typical than previously believed by psychologists and is even greater among gifted students than those of lesser intelligence.

4. The remediation of deficiencies and the enrichment of strengths is predicated upon a basic belief that all gifted children should be well-rounded in their cognitive development. In reality, we might well have to choose between having a flat SOI profile and a self-actualized child as educational objectives.

5. The SOI model has severe psychometric limitations. These center around the lack of reliable assignment of test items to SOI categories, unreliable difference scores, and the lack of adequate validity evidence for differential treatment effects.

6. More adequate normative data must be gathered in order to interpret properly purported strengths and weaknesses.

7. The absence of adequately described normative data and data on the standard errors of measurement of the profiles constitutes a risky practice when the standard errors of measurement preclude ethical and psychometrically sound use of the SOI model.

8. Much of the limited evidence indicating that SOI abilities can be trained and enhance school achievement suffers from a variety of methodological difficulties. These problems (absence of control groups, practice effects, regression toward the mean, nonspecific treatment effects, failure to compare SOI teaching approaches with other instructional approaches) render the evidence equivocal.

9. Given the limited and frequently unsupportive nature of available research evidence, SOI analyses are better reserved for research purposes than clinical purposes. Counselors, or other educators, involved in gifted programs should not look upon the SOI model as one whose practical usefulness has been established.

10. Counselors and other educators interested in the identification and education of the gifted must guard against the tendency to use technically inadequate tests when more valid instruments are available. Insisting that SOI based instruments and analyses meet the minimum standards for reliability and validity will go far toward promoting their ethical and suitable usage.

REFERENCES

American Psychological Association/American Educational Research Association/National Council on Measurement and Evaluation (1974). Standards for educational and psychological tests. Washington, DC: American Psychological Association.

Ausubel, D., & Sullivan, E. (1970). Theory and problems of child development. New York: Grune & Stratton.

Cattell, R. B. (1971). Abilities: Their structure, growth and action. Boston: Houghton Mifflin.

Clarizio, H., & Bernard, R. (1981). Recategorized WISC-R scores of learning disabled children and differential diagnosis. Psychology in the Schools, 18, 5-12.

Clarizio, H., & Veres, V. (1984). WISC-R patterns of emotionally impaired and diagnostic utility. Psychology in the Schools, 20, 409-414.

Cronbach, L. J. (1970). Essentials of psychological testing. New York: Harper and Row.

Dyer, C., Neigler, C., & Milholland, J. (1975). Rater agreements in assigning Stanford-Binet items to Guilford's structure of intellect operations categories. Journal of School Psychology, 13, 114-118.

Eysenck, H. J. (1967). Intelligence assessment: A theoretical and experimental approach. British Journal of Educational Psychology, 37, 81-98.

Eysenck, H. J. (1973). The inequality of man. London: Temple Smith.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Guilliksen, H. O. (1950). Intrinsic validity. American Psychologist, 5, 511-517.

Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. Psychological Bulletin, 80, 33-43.

Jensen, A. R. (1980). Bias in mental testing. New York: Free Press.

Kaufman, A. (1981). The WISC-R and learning disabilities assessment: State of the art. Journal of Learning Disabilities, 14, 520-526.

Kester, E. (1982). SOI: A qualitatively different program for the gifted. Gifted Child Quarterly, 26, 21-25.

McNemar, Q. (1964). Lost: Our intelligence? Why? American Psychologist, 19, 871-882.

Meeker, M. N. (no date). Using cognitive abilities as the diagnostic basis for curriculum planning. El Segundo, CA: Structure of Intelligence Institute.

Meeker, M. N. (1969). The structure of intellect. Columbus, OH: Charles Merrill.

Meeker, M. N. (copyright 1979, revised 1981, April). Using SOI test results: A teacher's guide. El Segundo, CA: Structure of Intelligence Institute.

Millichamp, D. A. (no date). A factor analysis of SOI-Binet patterns of bright four-year-old children. In Structure of Intelligence Institute, Collected readings: Vol. I. El Segundo, CA: Structure of Intelligence Institute.

O'Tuel, F. S., Ward, M., & Rawl, R. K. (1983). The SOI as an identification tool for the gifted: Windfall or washout? Gifted Child Quarterly, 27(3), 126-134.

Pearce, N. (1983). A comparison of the WISC-R, Raven Standard Progressive Matrices, and Meeker's SOI-Screening Form for Gifted, Gifted Child Quarterly, 27, 13-19.

Sattler, J. (1982). Assessment of children's intelligence and special abilities. Boston: Allyn & Bacon.

Silverstein, A. B. (1982). Pattern analysis as simultaneous statistical inference. Journal of Consulting and Clinical Psychology, 50, 234-240.

Stenson, C. M. (1982). Note on concurrent validity of structure of intellect gifted screener with Wechsler Intelligence Scale for Children-Revised. Psychological Reports, 50, 552.

Sternberg, R. (1979). The nature of mental abilities. American Psychologist, 34, 214-230.

Structure of Intelligence Institute (no date). Collected readings: Vol. I. El Segundo, CA: Author.

Structure of Intelligence Institute (no date). SOI research studies. El Segundo, CA: Author.

Structure of Intelligence Institute (1981). Technical data manual. El Segundo, CA: Author.

Thurstone, L. (1938). Primary mental abilities. Psychometric Monographs (No. 1).

Vernon, P. E. (1965). Ability factors and environmental influences. American Psychologist, 20, 723-733.

Vernon, P. E. (1979). Intelligence: Heredity and environment. San Francisco: W. H. Freeman and Company.